



Economía aplicada al estudio de la evolución de un curso*

Maria J. Blesa, Amalia Duch, Joaquim Gabarró, Jordi Petit, Maria Serna

Departament de Ciències de la Computació
Universitat Politècnica de Catalunya – BarcelonaTech,
08034 Barcelona

mjblesa@cs.upc.edu, duch@cs.upc.edu, gabarro@cs.upc.edu, jpetit@cs.upc.edu, mjserna@cs.upc.edu

Resumen

En este trabajo se propone un método para analizar cuantitativamente la efectividad de las diferentes medidas adoptadas para mejorar el proceso de evaluación continua de una asignatura. Nuestra propuesta utiliza herramientas simples provenientes de la economía y de las ciencias sociales. Primero se analiza el efecto de dichas medidas sobre las categorías de calificaciones obtenidas por los estudiantes en relación al incremento de carga docente del profesorado. En el estudio de la evolución de las calificaciones y el trabajo de los profesores a lo largo del tiempo se utiliza una aproximación coste-beneficio marginal. En segundo lugar se realiza un análisis de desigualdad utilizando distintos tipos de desagregación de datos y en varias subpoblaciones relacionadas con las calificaciones de los estudiantes. Finalmente se analiza la evolución de indicadores estadísticos como la *media* y la *satisfacción*. Dicho método se aplica al estudio de la asignatura de Programación-1 de la Facultat d'Informàtica de Barcelona de la Universitat Politècnica de Catalunya en los primeros 5 cursos de implantación del Grado en Ingeniería Informática. La metodología propuesta pretende introducir nuevas técnicas que permitan un análisis objetivo del impacto de medidas docentes en cualquier asignatura universitaria.

Palabras clave: Programación-1, evaluación continua, productividad, desigualdad, satisfacción, coeficiente de Gini.

Recibido: 17 de noviembre de 2016; **Aceptado:** 12 de diciembre de 2016.

1. Introducción

En Economía y en otras áreas del conocimiento es usual desde hace mucho tiempo el medir —entre otras cosas— diferentes cuestiones relacionadas con la utilización de recursos, la productividad, los costes de realizar algunas tareas, lo eficiente o eficazmente que se realizan, etc. En el ámbito educativo es menos usual preocuparse por estos aspectos. Todos los docentes sabemos que los recursos son limitados y que en los últimos tiempos de crisis estos se han reducido aún más. Sin embargo, no se suele medir de manera cuantitativa cuan eficientes, productivas, o eficaces están resultando algunas prácticas docentes.

En nuestra opinión es innegable que en los tiempos que corren es necesario y útil el preocuparse por estas cuestiones en el ámbito docente. Necesario porque como ya hemos men-

cionado los recursos son limitados y no nos podemos permitir el lujo de malgastarlos. Útil porque el poder contar con información objetiva y cuantitativa puede ayudar a una mejor planificación y a la toma de decisiones. Por lo expuesto, en este trabajo proponemos (y mostramos) que es posible (y sencillo) utilizar en el ámbito educativo, específicamente en la docencia universitaria, algunos de los instrumentos desarrollados por los economistas para establecer objetivamente el efecto de los cambios que se hacen en las asignaturas. Para ello, analizamos a fondo la asignatura Programación-1.

Programación-1 (PRO1) es la primera asignatura de programación del Grado de Ingeniería Informática en la Facultat d'Informàtica de Barcelona (FIB) de la Universitat Politècnica de Catalunya (UPC). Desde septiembre de 2006 el curso sigue una metodología «*learning-by-doing*» (aprendizaje mediante práctica) [1, 5, 19, 21] basada en una colección estraté-

*Una versión preliminar y reducida de este trabajo fue presentado en las XXII Jornadas sobre la Enseñanza Universitaria de la Informática (JENUI), que tuvieron lugar en Almería en julio de 2016.

gicamente seleccionada y organizada de ejercicios de programación que deben ser resueltos usando C++ con un juez de programación. Desafortunadamente, los dos primeros años de implantación no dieron los resultados esperados en la asignatura [10].

Por este motivo, el profesorado de PRO1 ha introducido, paulatinamente, una serie de medidas con el propósito de motivar a los estudiantes a trabajar más, mejor y de manera más autónoma y continua. Todo ello sin sacrificar los objetivos generales, el nivel requerido y el enfoque del curso. En consecuencia, el curso ha experimentado diversas modificaciones que han incrementado simultáneamente el trabajo de evaluación continua de los estudiantes y la carga docente del profesorado de la asignatura.

El análisis de la evolución temporal de las actividades evaluadoras de un curso es un tema de estudio consolidado [16, 23]. Con el fin de analizar objetivamente las diferentes medidas pedagógicas ya introducidas en la asignatura de PRO1 y, sobre todo, de poder predecir el efecto de nuevas medidas antes de su introducción, en un trabajo anterior se definió un indicador de la productividad en el aprendizaje [4]. Con este indicador se correlacionaron los resultados académicos obtenidos por los estudiantes con el tiempo dedicado por los profesores de la asignatura. Basándose en la aproximación de coste-beneficio marginal clásica en economía [6, 13, 27], este estudio anterior permite relacionar el número de estudiantes que aprueban con la carga docente mediante una curva de productividad, de tal forma que es posible evaluar el impacto de las medidas pedagógicas tomadas.

En este trabajo refinamos y ampliamos el análisis publicado anteriormente en relación a la evolución de PRO1, desde el curso 2010–2011. Utilizamos dos criterios de separación o *desagregación* de datos. En primer lugar, separando el total de notas en los distintos rangos de nuestro sistema de calificaciones (Sobresaliente, Notable, etc.). Basándonos en la metodología propuesta en el trabajo precedente [4], aplicamos un análisis coste-beneficio por rangos de notas.

En segundo lugar, para llevar a cabo un estudio de desigualdad, reclasificamos las notas en otras categorías. En concreto desagregamos las calificaciones en base a un criterio que es habitual en sociología y es también utilizada por Thomas Piketty [17, Capítulo 7]. Esta desagregación se basa en las *tablas sociales* y la *aritmética política* de Denis Diderot del siglo XVIII [8]. En este trabajo hemos optado por desagregar las calificaciones de PRO1 en tres segmentos de acuerdo con la denominada distribución (10%, 40%, 50%). Así se analizan diferentes parámetros desagregando la población en: el 10% con mayor nota (*clase alta*), el 40% que sigue (*clase media*) y el 50% restante (*clase baja*). Para cada una de estos segmentos de la población aplicamos un análisis basado en la nota media, combinada con una medida de la desigualdad social utilizando el *coeficiente de Gini* [7, 12].

Finalmente, además de dicho coeficiente, se toman en consideración otros indicadores como son la *media*, la *mediana*, la *desviación estándar* y la *satisfacción*. La satisfacción fue introducida por Amartya Sen [20] (Premio Nobel de Econo-

mía en 1998). La satisfacción permite tratar de modo conjunto la media y la desigualdad. De hecho se puede ver como una medida en que la media se pondera con el coeficiente de Gini para tener en cuenta la desigualdad. Concluimos el estudio con un análisis de la evolución de la media y de la satisfacción para cada una de las categorías consideradas.

La complejidad de un curso como PRO1 no puede captarse en su totalidad mediante un análisis cuantitativo a partir de los datos de que disponemos. Sin embargo, los autores sí creemos que los resultados obtenidos y la metodología propuesta pueden ser de gran utilidad al considerar la introducción de mejoras tendentes a conseguir un mayor número de estudiantes que alcanza el nivel requerido en la asignatura de programación.

2. Contexto

La FIB ofrece el curso de PRO1 dos veces en cada año académico: en el semestre de otoño (S1) y en el de primavera (S2). En S1 llegan todos los estudiantes de nueva admisión y, para un gran número de ellos, PRO1 es su primer curso de programación. En S2 los estudiantes son mayoritariamente repetidores. A pesar de la diferente naturaleza de los estudiantes de cada semestre, ni los contenidos del curso ni la forma de evaluación cambian entre S1 y S2. Lo que sí se suele adaptar a la audiencia concreta de cada semestre es la organización de las clases, especialmente el de las clases prácticas. La tendencia conjunta de ambos semestres (y por tanto del curso) es la que se ve reflejada en el S2 [4] que es el que consideraremos en este análisis.

El curso de PRO1 se compone de 2 horas de clases de teoría y 3 horas de clases de práctica a la semana. En ambos semestres, las clases de teoría se organizan en grupos de 60–80 estudiantes y las de práctica en grupos de 15–20 estudiantes.

El objetivo principal de PRO1 siempre ha sido asegurar que los estudiantes aprenden y dominan habilidades básicas de programación. Los contenidos del curso se muestran en el cuadro 1. A partir del curso 2006–2007 se hizo una apuesta por alcanzar dicho objetivo con un claro refuerzo de la parte práctica de la asignatura. Por eso el curso se reorganizó alrededor de concepto de «problemas de programación», es decir, ejercicios de programación descritos con enunciados muy claros en términos de entradas válidas y salidas correctas. Para cada ejercicio, los estudiantes han de escribir un programa en C++ que lo resuelva y se comporte acorde con el enunciado. Para aplicar esta metodología de aprendizaje mediante práctica, se hace uso del *juez de programación* Jutge.org [11]. El juez es una herramienta en línea que, automáticamente y en tiempo real, verifica si el programa propuesto por un estudiante es una solución correcta a un ejercicio de programación y en algunos casos otorga *awards* reconociendo actuaciones singulares con el fin de motivar mejor el trabajo. Dicha herramienta de soporte está disponible para los estudiantes las 24 horas del día, proporciona una amplia fuente de ejercicios y fomenta la autoorganización del proceso de aprendizaje. A nivel docente

1. Introducción a los principios básicos de la programación.
Introducción de conceptos fundamentales: Algoritmo, programa, variable, expresión, tipos de datos.
Instrucciones básicas del lenguaje C++.
2. Instrucciones iterativas, Ejemplos.
3. Esquemas de recorrido y búsqueda.
Secuencias. Recorridos y búsquedas en secuencias.
4. Acciones y Funciones.
Paso de parámetros. Ámbitos de visibilidad.
5. Recursividad. Introducción al diseño recursivo.
6. Tablas. Tablas unidimensionales. Tablas multidimensionales. Recorridos y búsquedas en tablas.
7. Agregaciones de datos.
8. Algoritmos fundamentales I. Algoritmos de ordenación. Búsqueda binaria.
9. Algoritmos fundamentales II. Otros algoritmos fundamentales: Horner, producto rápido, exponenciación rápida, etc.

Cuadro 1: Contenidos de PRO1

permite dinamizar enormemente las clases prácticas y hacer un fácil seguimiento del trabajo de los estudiantes, así como diseñar, organizar y administrar exámenes. Este tipo de jueces surgieron de concursos de programación (por ejemplo, el UVa Online Judge [18]) y su uso ha sido adaptado en los últimos tiempos a entornos educativos [11, 14, 26, 28]. La figura 1 muestra el tablero de un usuario de Judge.org y un ejemplo de sus enunciados.

La primera colección de ejercicios se diseñó convenientemente en 2006 y se organizó por temas y dificultad. El tamaño actual de la colección, con más de 300 problemas, permite tener material suficiente tanto para la selección de los contenidos de las sesiones de laboratorio, como para ejercicios complementarios. Durante el curso se espera que los estudiantes resuelvan el máximo número de problemas posible. En el curso 2006–2007 los exámenes también se adaptaron a este formato y se conformaron de problemas de dificultad similar a los de la colección.

Evolución. A lo largo de los cursos se han ido tomando una serie de medidas para intentar incrementar el número de aprobados sin merma de contenidos ni conocimientos requeridos. Siguiendo la notación introducida en el trabajo precedente [4], hacemos referencia a cada una de las medidas introducidas desde el curso 2010–2011 al curso 2014–2015 como t_4, \dots, t_8 , respectivamente. Haciendo abuso de notación, usaremos cada uno de los términos t_i para denotar tanto las medidas introducidas en el curso correspondiente, como para denotar el propio curso. Comenzamos nuestro análisis con el curso 2010–2011 por corresponder al curso de puesta en marcha del Grado en Ingeniería Informática.

t_4 Grado en Informática (2010–2011): Se introduce un nuevo plan de estudios, que sigue vigente actualmente. El

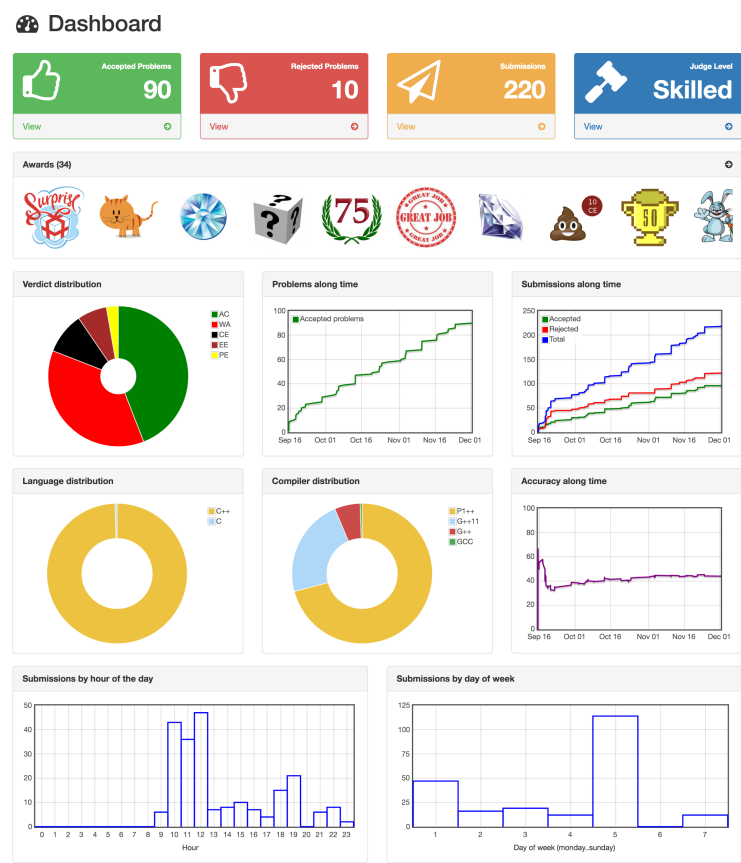
curso se estructura en 2 horas de teoría y 3 horas de laboratorio semanales. La evaluación continuada incluye la resolución de ejercicios prácticos en forma de exámenes parciales con un peso importante (60%) en la nota final.

t_5 Listas de problemas (2011–2012): Para fomentar el trabajo continuado y facilitar una evaluación continua positiva, se diseñan listas de problemas adicionales para cada tema del curso. El derecho a realizar un examen parcial se condiciona a resolver como mínimo el 70% de los ejercicios de las listas correspondientes. En contrapartida los parciales se componen de ejercicios de dichas listas.

t_6 Curso de reevaluación (2012–2013): La FIB ofrece la posibilidad de realizar un curso intensivo y un examen adicional de reevaluación a los estudiantes suspendidos con nota igual o superior a 3,5. Dicho curso es una síntesis de 12 horas y tiene lugar una vez ha acabado el curso convencional. La asistencia a las clases es obligatoria y se mantiene el requisito de resolver como mínimo el 70% de los ejercicios propuestos para tener derecho a examen. Para aquellos alumnos que aprueban el examen de reevaluación la nota final es 5,0; el resto mantiene la calificación del curso.

t_7 Exámenes parciales con nuevos ejercicios (2013–2014): Se cambia el diseño de los exámenes parciales. Ahora pasan a componerse exclusivamente de problemas nuevos que los estudiantes no conocen de antemano. Se pretende evitar un problema de mecanización del aprendizaje.

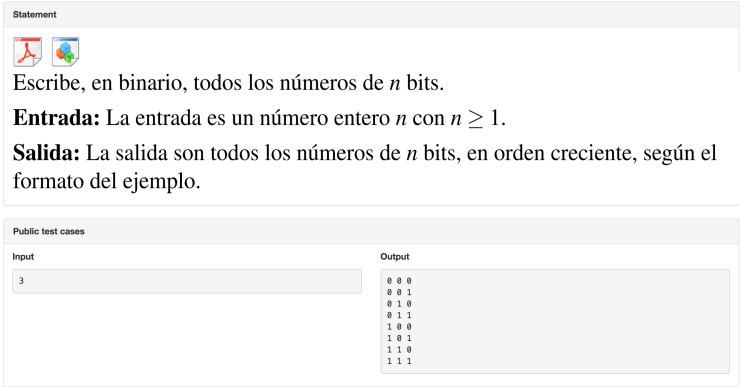
t_8 Diversificación del curso (2014–2015): Hasta 2014 la colección de problemas seleccionada para PRO1 se man-



(a) Tablero de un usuario

Ceros y unos

P20867 es



(b) Ejemplo de problema

Figura 1: Capturas de pantalla de Jutge.org.

tuvo prácticamente invariada. Con el objetivo de diversificar los contenidos del curso y ayudar a los estudiantes a afrontar y resolver problemas diferentes, se incorporan problemas nuevos a la colección y se diseñan colecciones distintas para los semestres S1 y S2.

Obsérvese que estas medidas no son sustitutivas, sino que se acumulan en el tiempo; es decir, las medidas que se toman en un determinado t_i siempre incluyen las de t_j anteriores ($4 \leq j < i$). Esto nos permitirá tener una cierta visión de su evolución en el tiempo y de observar tendencias.

Trabajo. La figura 2 muestra, para cada periodo, la carga estimada de trabajo total de los docentes (W) medido en horas y, asumiendo que el trabajo por matriculado es uniforme, la estimación del trabajo dedicado a los aprobados (W_{apr}). Referimos al lector al artículo anterior [4] para una descripción completa del desglose de la estimación del trabajo y de la metodología utilizada para su obtención. La estimación de la cantidad de trabajo se obtuvo mediante una encuesta entre el profesorado de la asignatura (aproximadamente, unos 20 docentes) que incluía una valoración de las horas destinadas a cada una de las siguientes tareas: (1) Diseño, prueba y preparación de exámenes; (2) sesiones de teoría; (3) sesiones de laboratorio; (4) corrección de exámenes; (5) vigilancia de controles; (6) coordinación del curso; (7) rediseño del curso; y (8) mantenimiento del software.

Calificaciones. Centramos este estudio en el análisis de las calificaciones obtenidas en el S1 de cada periodo. El comportamiento del S2 es muy distinto al del S1 [4] al tratarse de alumnos que en su mayor parte son repetidores. Sin embargo, los resultados para los datos agregados del S1 y S2 no muestran diferencias significativas con las del S1.

3. Productividad

El objetivo de este apartado es extender el análisis coste-beneficio hecho anteriormente [4] a los grupos correspondientes a una desagregación clásica de las notas. Queremos analizar cómo se desagrega la función de productividad y estudiar las variaciones en cada grupo de estudiantes. Es decir, queremos determinar si el efecto (positivo o negativo) de una medida con respecto a la productividad del total de aprobados se mantiene o cambia dependiendo de la categoría de notas.

Categorías de notas. Para realizar el estudio de la productividad por grupos hemos desagregado las calificaciones (en el rango NP , $[0, 10]$) en seis rangos. Tres de ellos corresponden a los aprobados Apr ($[5, 10]$) y se corresponden con la calificación habitual: *Sob*, sobresaliente ($[9, 10]$); *Not*, notable ($[7, 9]$); *Suf*, suficiente ($[5, 7]$). Las notas de los no aprobados Sus se han desagregado en 3 rangos: *Rev*, reevaluable ($[3.5, 5]$); *NR*, no reevaluable ($[0, 3.5]$); *NP*, no presentado. Entendemos que el rango de reevaluables es importante ya que algunas de

las medidas tomadas pretenden incidir específicamente en este grupo. En nuestros estudios la calificación de NP se otorga cuando el estudiante no ha participado en ningún acto evaluativo con peso superior a 20 %. Tales actos evaluativos tienen lugar al final del curso por lo que no podemos presuponer un abandono temprano de la asignatura. En el cuadro 2 se muestra el desglose de las calificaciones para las notas del S1 según las categorías propuestas. *Est* representa el número total de estudiantes matriculados.

Productividad. La productividad Π de un curso se define como [4]

$$\Pi = \frac{P}{W},$$

donde P es el número de estudiantes aprobados (categoría *Apr*) y W el número de horas de dedicación docente realizado por parte del profesorado. En este trabajo utilizamos esta medida de productividad aplicada a cada una de las distintas poblaciones, es decir, que en vez de utilizar el total de aprobados en la fórmula anterior, utilizamos el número de estudiantes que se encuentra en cada una de las categorías de estudio. Así

$$\Pi(C) = \frac{\#C}{W}$$

donde $\#C$ es el número de estudiantes en la categoría C objeto de estudio. Notemos que, si $n = \#Est$ es el número de estudiantes, el trabajo docente por estudiante es $W_{ind} = W/n$. Así la productividad en la categoría *Est* es

$$\Pi(Est) = \frac{\#Est}{W} = \frac{n}{W} = \frac{1}{W_{ind}},$$

es decir la inversa del trabajo docente dedicado a cada estudiante.

La noción de productividad útil refina el análisis de las diferentes categorías de notas dentro de la categoría de estudiantes aprobados. La definimos como

$$\Pi_u(C) = \frac{\#C}{W_{apr}}$$

siendo W_{apr} la estimación del trabajo total dedicado a los aprobados, es decir, $W_{apr} = \#Apr \cdot W_{ind}$.

Obsérvese que un incremento de la productividad significa que se obtienen los mismos resultados con menos horas de dedicación de los profesores; no significa, por lo tanto, que se obtienen mejores resultados con el mismo trabajo. En este último caso se pasaría a otra categoría y se reflejaría en la productividad útil.

Análisis de la productividad. Las figuras 3a, 3b y 3c muestran las curvas de productividad por categoría correspondientes a las calificaciones del S1. Pese a no representar medidas continuas en el tiempo, dichas gráficas se muestran en formato continuo para facilitar la legibilidad simultánea de los muchos datos representados, y para dar una idea de tendencia. Como mencionamos anteriormente, las medidas t_i no son sustitutivas, sino que se acumulan en el tiempo.

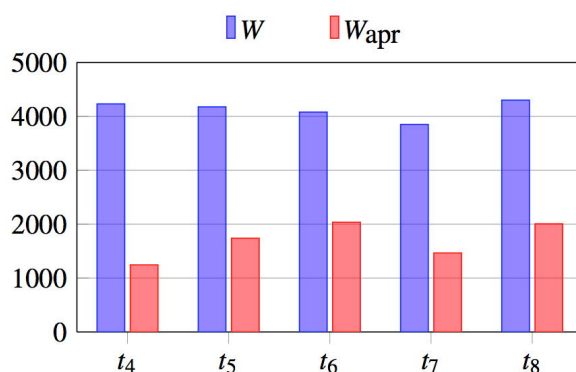


Figura 2: Trabajo total y trabajo destinado a *Apr*.

	Est	Apr	Sob	Not	Suf	Sus	Rev	NR	NP
t ₄	514	132	11	36	85	382	65	243	74
t ₅	492	191	31	64	96	301	56	198	47
t ₆	465	211	44	86	81	254	57	127	70
t ₇	436	155	11	49	95	281	36	152	93
t ₈	448	209	24	72	113	239	32	101	106

Cuadro 2: Número de estudiantes en cada rango de notas en los cursos desde t_4 a t_8 . Como procede, $Est = Apr + Sus$, $Apr = Sob + Not + Suf$ y $Sus = Rev + NR + NP$.

En las figuras 3a, 3b y 3c puede verse que la curva de productividad $\Pi(Est)$, aún con poca variabilidad, es decreciente, lo que indica un incremento continuado del trabajo docente por estudiante. Con respecto la productividad en los distintos grupos, cabe mencionar que los grupos *Sob* y *Not* muestran el mismo comportamiento. La tendencia es similar a la curva de productividad de *Apr*. Esto permite observar que las medidas adoptadas en la asignatura tienen el mismo efecto en la productividad relativa a estas dos poblaciones y que las diferencias en tendencia con la curva de productividad de *Apr* provienen de la productividad de la categoría *Suf*. Como se puede observar, la gráfica de *Suf* no tiene la misma forma que la de *Apr*, y en t_6 la productividad en *Suf* baja mientras que no lo hace en *Sob*, *Not* y *Apr*. Esto indica que la medida tomada en t_6 tiene un efecto contrario en *Suf* que en *Sob* y *Not*.

Las gráficas de *NR* y *Apr* tienen tendencias cuasi simétricas (o complementarias) lo que parece indicar que la variación de productividad de *Apr* contrarresta, en cierta medida, la variación en productividad de *NR*.

Parece muy notorio que el efecto de todas las medidas adoptadas es contrario en la categoría *Apr* y en la categoría *NR*, ya que las curvas son cuasi simétricas y en ambos casos con pendientes bastante pronunciadas. Esto indicaría que una productividad creciente de aprobados se corresponde con una productividad decreciente de no reevaluables, y viceversa. Por tanto, las medidas afectan en sentido inverso a las dos categorías.

La curva de productividad de *Rev* es claramente decreciente, y decrece más rápidamente a partir de t_6 , lo que parece razonable al haberse introducido medidas orientadas a minimizar el número de estudiantes en esta categoría.

Con respecto a la categoría *NP* puede observarse que, excepto de t_4 a t_5 , la productividad es creciente y, de t_5 a t_7 , el aumento es bastante pronunciado. Sin embargo, la productividad real en esta categoría puede no ser visible al haber adoptado un criterio de reparto equitativo del trabajo docente con relación a los estudiantes matriculados.

Como primera conclusión vemos que el efecto en la productividad de las medidas que se han tomado no se desagrega con la misma tendencia (véanse las tendencias de la categoría *Apr* en las categorías *Sob* y *Not*). Con relación a las categorías de alumnos que suspenden, las tendencias de las categorías *Rev* y *NR* son similares, con una tendencia diferente en la categoría *NP*.

Análisis de la productividad útil. La figura 3d muestra las curvas de productividad útil para las poblaciones *Sob*, *Not* y *Suf*. El resultado, reafirma la observaciones sobre la curva de productividad, con un efecto más pronunciado. La productividad en *Suf* parece ir en sentido contrario que en *Sob* y *Not*. Esto refleja que las horas de trabajo W_{apr} son productivas principalmente en las categorías de notas superiores.

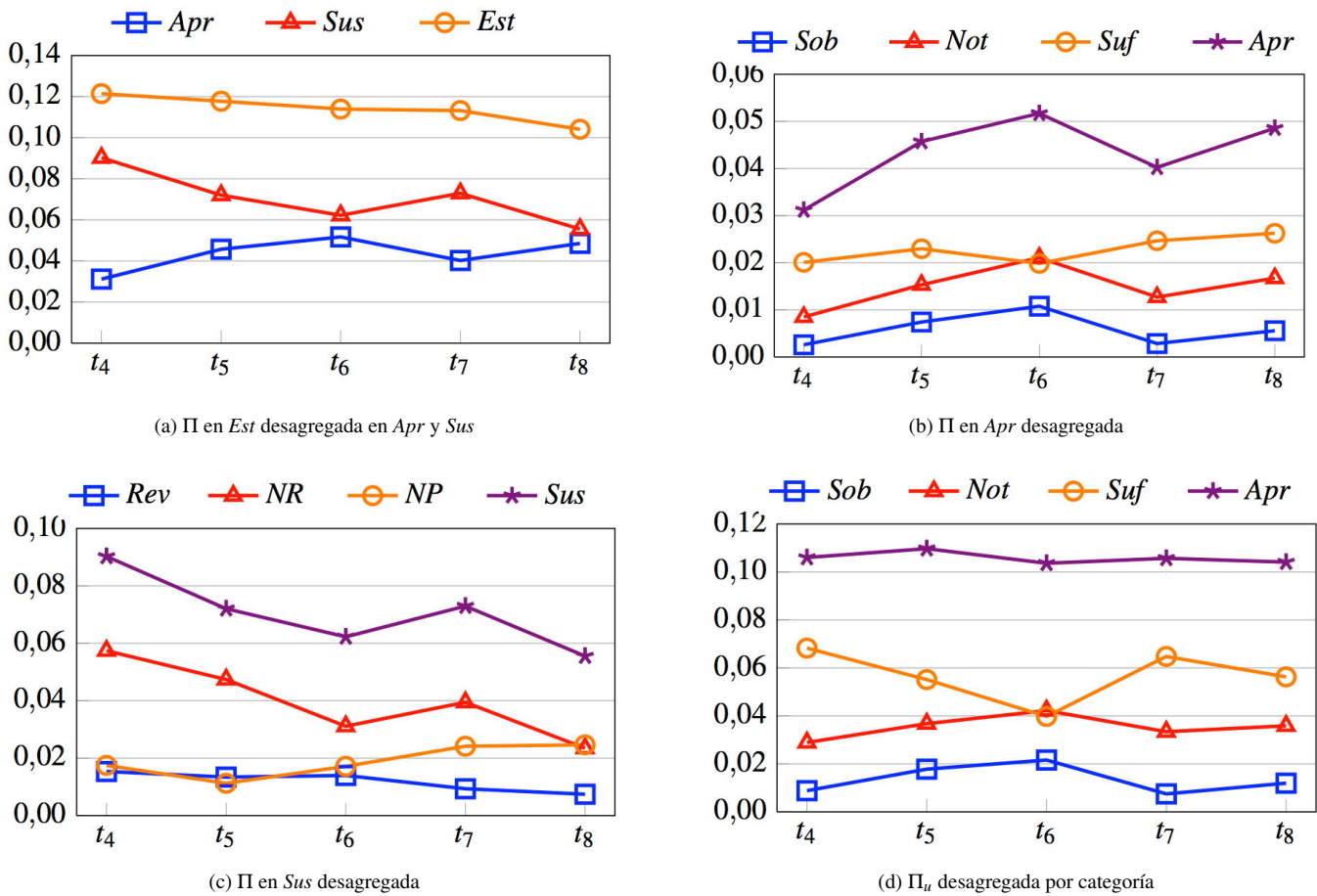


Figura 3: Desagregaciones de la productividad.

	$\Delta W_t > 0$	$\Delta W_t < 0$
$\Delta C_t < 0$	(NEG+) Muy negativa	(NEG) Negativa
$\Delta C_t > 0$	$\Delta W_t > \Delta C_t$ (EVI) A evitar $\Delta W_t < \Delta C_t$ (POS) Positiva	(POS+) Muy positiva

Cuadro 3: Clasificación de medidas [4]

Ganancia marginal. La noción de ganancia marginal [4], complementa el análisis de la productividad proporcionando una mayor exactitud en el análisis de las variaciones entre periodos consecutivos. Extendiendo el modelo a la población desagregada definimos la variación de trabajo como

$$\Delta W_t = \frac{W_t}{N_t} - \frac{W_{t-1}}{N_{t-1}},$$

donde W_t es el trabajo total en el período t y N_t el número total de matriculados. Definimos también la variación en la categoría C como

$$\Delta C_t = \frac{C_t}{N_t} - \frac{C_{t-1}}{N_{t-1}},$$

donde C_t es el número de estudiantes en la categoría C en el período t . Usando los valores mostrados en el trabajo anterior [4], se clasifican las medidas en cinco categorías, tal y como se refleja en el cuadro 3, y se asocia un nivel de bondad a la medida implementada. El cuadro 4 proporciona los valores numéricos de la variación para las categorías *Sob*, *Not*, *Suf* y *Apr* junto con la matriz de clasificación alfabética. Como se puede apreciar, la medida introducida en t_5 es muy positiva y lo es tanto en la categoría *Apr* como en todas sus subcategorías. Lo mismo ocurre con t_8 para una medida que se puede clasificar como positiva. El efecto de t_6 y t_7 es opuesto para la categoría *Suf*. Así, t_6 es muy negativa para *Suf* y positiva en el resto mientras que t_7 es positiva para *Suf* y muy negativa para el resto. Se reafirma así la tendencia opuesta que se apreciaba en los análisis de productividad.

En este análisis de productividad tratamos la nota *NP* como si fuese un cero, ya que solo podemos tratar esta categoría de estudiantes como estudiantes activos a los que, en consecuencia, se les asignan recursos. Dado que los estudiantes abandonan a lo largo del curso y lo hacen de múltiples maneras sería necesario un estudio adicional para poder reevaluar el trabajo total asociado a estos estudiantes. Por este motivo no hemos analizado la productividad sobre la subcategoría de estudiantes presentados (aquellos que participan en los actos evaluativos con mayor peso). En el resto del estudio sí que analizaremos los resultados para este conjunto de estudiantes.

4. Desigualdad

Hasta ahora se han analizado los resultados desagregando las notas según la división académica usual. Esta sección es un primer intento de análisis desagregando las notas utilizando algunos de los métodos bien conocidos en ciencias sociales y economía. También se pretende ver hasta qué punto dichos métodos permiten arrojar nueva luz a los análisis anteriores o permiten mostrar aspectos complementarios. En consecuencia, proponemos un análisis transversal que se basa en la evolución de dos componentes: Por una parte la nota media y la nota de corte de los segmentos y por otra el coeficiente de Gini [7, 12].

Segmentos de la población y nota media. La distribución (10%, 40%, 50%) con relación a una medida de riqueza (por ejemplo la renta) consiste en desagregar la población en tres sectores [17]. El segmento alto formado por la población que corresponde al 10% con valor más alto. El segmento medio, correspondiente al 40% siguiente. El 50% restante de la población constituye el segmento bajo. En nuestro símil, utilizaremos las calificaciones como medida de riqueza de la población de estudiantes. Usaremos los identificadores *Ea*, *Em* y *Eb* para referirnos a los sectores alto, medio o bajo de la población de estudiantes matriculados, *Est*. De la misma manera *Pa*, *Pm*, *Pb* identifican dichos sectores en la subpoblación de estudiantes presentados, *EstPr*. Finalmente, *Aa*, *Am*, *Ab* identifican los sectores de la población de estudiantes aprobados, *Apr*. En el cuadro 5 se muestran las evolución de las notas medias y notas de corte en cada uno de los sectores. Como era de esperar los valores son ligeramente mejores en *EstPr* que en *Est* mientras que en *Apr* los valores mejoran notablemente.

Con este tipo de desagregación podría intentarse hacer el mismo análisis de productividad que se hizo para la desagregación por notas, sin embargo, en este caso dicho análisis no es útil. Recordemos que la productividad para el segmento *S* es

$$\Pi(S) = \frac{1}{W_{ind}} \cdot \frac{\#S}{n}.$$

Cuando *S* es un porcentaje fijo de la población, su productividad es la inversa del trabajo individual escalada por el porcentaje correspondiente. Así, las curvas de productividad de la figura 3 no proporcionan información adicional. La alternativa que proponemos es estudiar la desigualdad mediante el coeficiente de Gini.

Coeficiente de Gini. La desigualdad de una sociedad se mide habitualmente mediante el coeficiente de Gini. Esta medida se ha utilizado en el análisis de indicadores económicos como el PIB [3, 17] pero también en ámbitos educativos [24, 25]. Formalmente, si $r = (r_1, \dots, r_n)$, con $n > 1$, es un vector de resultados, y su media es $\mu = \frac{1}{n} \sum_{i=1}^n r_i$, el coeficiente de Gini se define como

$$gini = \frac{\sum_{i=1}^n \sum_{j=i+1}^n |r_i - r_j|}{\mu \cdot n \cdot (n-1)}.$$

Cuando r representa un conjunto de notas, el coeficiente de Gini mide lo separadas que están unas de otras y, por lo tanto, mide el grado de heterogeneidad de las notas.

Intuitivamente, un coeficiente de Gini próximo a 1 indica una fuerte desigualdad, mientras que uno próximo a 0 indica una distribución igualitaria. Así pues, en nuestro contexto, el coeficiente de Gini mide el grado de heterogeneidad (o desigualdad) de conocimientos de los estudiantes. Parece tentador asociar un coeficiente de Gini alto a categorías con gran heterogeneidad de conocimientos. Un coeficiente alto sería un indicador de fuertes diferencias en cuanto a formación. El coeficiente de Gini del primer curso de programación podría servir como indicador del nivel de heterogeneidad para el siguiente curso.

	ΔSob	ΔNot	ΔSuf	ΔApr	ΔW
t_5	4,16	6,00	2,98	13,14	-0,09
t_6	3,16	5,49	-2,09	6,56	0,28
t_7	-6,94	-7,26	4,37	-9,83	0,06
t_8	2,83	4,83	3,43	11,10	0,77

(a) Categorías

	ΔSob	ΔNot	ΔSuf	ΔApr
t_5	POS+	POS+	POS+	POS+
t_6	POS	POS	NEG+	POS
t_7	NEG+	NEG+	POS	NEG+
t_8	POS	POS	POS	POS

(b) Trabajo

Cuadro 4: Variación en las medidas y en el trabajo.

	Est Med	Ea Med	$Ea-Em$ Corte	Em Med	$Em-Eb$ Corte	Eb Med
t_4	2,83	8,24	6,71	4,42	2,01	0,50
t_5	3,73	9,12	8,10	5,79	3,40	1,02
t_6	4,39	9,69	8,92	6,70	4,28	1,48
t_7	3,49	9,06	7,90	5,36	3,10	0,89
t_8	3,80	8,95	7,90	5,90	3,80	1,12

(a) Matriculados

	$EstPr$ Med	Pa Med	$Pa-Pm$ Corte	Pm Med	$Pm-Pb$ Corte	Pb Med
t_4	3,30	8,45	7,16	4,93	2,80	0,97
t_5	4,15	9,20	8,30	6,11	4,10	1,54
t_6	5,15	9,73	9,29	7,20	5,47	2,59
t_7	4,39	8,65	7,90	5,76	4,00	2,47
t_8	4,98	9,21	8,20	6,60	5,20	2,83

(b) Presentados

	Apr Med	Aa Med	$Aa-Am$ Corte	Am Med	$Am-Ab$ Corte	Ab Med
t_4	6,77	9,45	8,78	7,49	6,37	5,69
t_5	7,10	9,69	9,40	7,96	6,90	5,91
t_6	7,59	10,00	9,78	8,56	7,55	6,33
t_7	6,77	9,84	9,10	7,79	6,30	5,39
t_8	6,97	9,72	9,00	7,73	7,00	5,84

(c) Aprobados

Cuadro 5: Notas promedio y notas de corte para los distintos segmentos.

	<i>Est</i>	<i>Ea</i>	<i>Em</i>	<i>Eb</i>		<i>EstPr</i>	<i>Pa</i>	<i>Pm</i>	<i>Pb</i>		<i>Apr</i>	<i>Aa</i>	<i>Am</i>	<i>Ab</i>
t_4	0,54	0,06	0,18	0,66	t_4	0,47	0,05	0,14	0,50	t_4	0,11	0,02	0,06	0,04
t_5	0,47	0,03	0,13	0,57	t_5	0,41	0,03	0,10	0,46	t_5	0,12	0,01	0,05	0,05
t_6	0,43	0,02	0,11	0,55	t_6	0,33	0,01	0,09	0,35	t_6	0,11	0,01	0,04	0,06
t_7	0,49	0,04	0,14	0,62	t_7	0,36	0,03	0,10	0,45	t_7	0,14	0,01	0,05	0,04
t_8	0,46	0,04	0,11	0,63	t_8	0,29	0,04	0,08	0,29	t_8	0,11	0,02	0,04	0,05

(a) Coeficientes de Gini sobre matriculados

(b) Coeficientes de Gini sobre presentados

(c) Coeficientes de Gini sobre aprobados

Cuadro 6: Coeficientes de Gini

El cuadro 6 muestra la evolución temporal de los coeficientes de Gini para las poblaciones *Est* y *Apr*. Sobre *Est* los valores de los coeficientes se sitúan alrededor de 0,5, indicando poca variabilidad en el nivel de desigualdad y un tipo de desigualdad media. Sobre *EstPr* los valores de los coeficientes se sitúan entre 0,29 y 0,47, indicando una mayor variabilidad en el nivel de desigualdad. La situación es muy distinta sobre la población de aprobados. En este caso el coeficiente de Gini se sitúa alrededor de 0,1, indicando un nivel bajo de desigualdad.

Análisis. Con relación al total de estudiantes (cuadro 5a) vemos que casi siempre el 10 % alto tiene una nota próxima a 9. El 40 % central está aproximadamente comprendido entre 4,5 y 7. Para el 50 %, aunque a veces incluye notas más altas que 3, la media es inferior a 1,5. Como ya hemos comentado la nota media de *Eb* está influenciada por la asignación de nota 0 a los no presentados. Apreciamos dos hechos importantes. Primero, la nota media del 50 % inferiores es muy baja (por debajo del 1,5), esto muestra un fenómeno de cola larga (*long tail*) que parece apartarse del comportamiento gaussiano. En una primera lectura la cola larga se puede interpretar como una tasa de fracaso alta, aunque no hay que olvidar que se trata de un curso de fase selectiva. Segundo, notemos que la nota de corte entre el 40 % y el 50 % oscila en la mayor parte de los casos alrededor del 3,5. Este hecho permite dar una explicación “sociológica” *a posteriori* del mecanismo de reevaluación. Este mecanismo está pensado como un incentivo para los estudiantes que están a caballo entre *Em* y *Eb*. Esto permite dar una contrapartida numérica al hecho de que los estudiantes con derecho a la reevaluación son los que «potencialmente pueden aprobar». Es decir, se pretende incentivar a la clase media-baja y la baja-alta.

Cuando se trata sólo con presentados *EstPr* (cuadro 5b), el 50 % bajo oscila más alcanzando en un caso casi un 3 y en otro casi un 5,5. El 10 % alto tiene una nota casi siempre por encima del 8,5. Por otra parte el 40 % central presenta una cierta variabilidad. Esto parece indicar que la distribución (10 %, 40 %, 50 %) permite apreciar diferencias en las clases media y baja que no se aprecian en las otras subpoblaciones.

Cuando se trata sólo con aprobados *Apr* (cuadro 5b), el 50 % bajo no deja de estar alrededor del 6, aunque el 10 % alto tiene una nota alrededor del 10 y el 40 % central alrededor del 7. Las notas de corte en este caso parecen oscilar alrededor

de las categorías tradicionales. Esto parece querer decir que la distribución (10 %, 40 %, 50 %) y la distribución mediante notas convencionales capturan —aunque no de modo idéntico— el mismo tipo de fenómeno.

Analizamos ahora nuestros datos con el coeficiente de Gini. En el caso de todos los matriculados, el coeficiente de Gini es aproximadamente 0,5. Con carácter puramente orientativo comparemos con algunos datos de Piketty [17, Capítulo 7], donde para la renta europea el coeficiente de Gini es de 0,19 y para la propiedad del capital es de 0,58. En economía un coeficiente de 0,5 se asocia a una desigualdad alta. Como primera aproximación podemos suponer lo mismo aquí. Sin embargo no hay que olvidar que este fenómeno está influenciado por la cola larga del 50 % que hemos visto antes. Notemos que cuando sólo se tiene en cuenta el conjunto de los presentados o de los aprobados los valores bajan. Como cabía esperar la desigualdad disminuye. En el caso de los aprobados baja drásticamente a un máximo de 0,14.

Proponemos un análisis comparando la variación de la nota media con la variación del coeficiente de Gini. En este contexto un incremento de la nota media junto con un decremento del coeficiente de Gini indica que se ha conseguido incrementar la nota de forma más igualitaria (las notas estarán mas cercanas a la media) mientras que si el coeficiente de Gini aumenta se detecta una mayor desigualdad en el reparto de las calificaciones.

Bajo este criterio se puede observar que, tanto en *Est* como en *Apr*, t_7 es la única medida que hace bajar la media y con la que aumenta la medida de la desigualdad en las dos categorías. Reafirmandose con este análisis la percepción, ya apuntada por el análisis coste-beneficio, de que la medida t_7 ha sido negativa. La medida t_8 contrarresta hasta cierto punto la medida previa sin llegar a alcanzar los niveles de t_6 .

5. Indicadores adicionales

En esta sección estudiamos la evolución de algunas medidas adicionales de agregación de datos, complementando así el análisis del coeficiente de Gini. La esperanza (o media), $\mu(r)$, que proporciona la medida más conocida de agregación de los valores de una población r , por ejemplo cuando hablamos de *renta per cápita*. Por otra parte, el coeficiente de Gini, $gini(r)$, que como hemos visto se utiliza para medir el

grado de heterogeneidad dentro de una población (por ejemplo en la distribución de la renta), nos permite considerar otra medida de agregación: la satisfacción, *sat*. Formalmente, si $r = (r_1, \dots, r_n)$, con $n > 1$, es un vector de resultados, y su media es $\mu = \frac{1}{n} \sum_{i=1}^n r_i$, la satisfacción [20] se define como

$$sat(r) = \mu(r) \cdot (1 - gini(r)).$$

Es posible ver la satisfacción como una versión matizada de la desigualdad de la media. Por ejemplo, en un país con una *renta per cápita* muy alta y con un reparto muy desigual tendremos un coeficiente Gini cercano a 1; intuitivamente, esto se podría interpretar como una baja satisfacción de las personas, y se reflejaría con el indicador $sat = \mu \times (1 - gini)$ cercano a 0. Posiblemente lo mismo se podría decir de las calificaciones en el entorno de la evaluación de un curso: si la media de nota es alta pero el reparto de las notas es muy desigual, la percepción del nivel central de la nota se desplaza hacia abajo y puede acercarse más a *sat* que a μ .

También calculamos otros estadísticos que permiten medir otros aspectos de los datos cuya consideración nos parece relevante para completar nuestro estudio. La mediana, *md*, que representa la nota de corte entre el 50 % de la población con menos nota y el otro 50 %. La desviación estándar, σ , que permite medir la concentración alrededor de μ .

En el cuadro 7 presentamos el resumen de los seis indicadores, en los cursos considerados, tanto para el conjunto de notas de todos los estudiantes matriculados (*Est*) como para el de los estudiantes presentados (*EstPr*). En el histórico de PRO1, tal y como cabe esperar del nivel de desigualdad, constatamos que μ y *sat* están alejadas por un factor cercano al 50 %. Este factor mejora considerablemente en algunos cursos entre los estudiantes que no abandonan la asignatura llegando a factores cercanos al 65–70 %. Con relación a la variación entre indicadores podemos observar que en las notas de matriculados se presentan únicamente dos tipos de comportamiento:

1. μ, σ y *sat* aumentan mientras que *gini* disminuye, o
2. μ, σ y *sat* disminuyen mientras que *gini* aumenta.

En este contexto estadístico una medida se podría clasificar como positiva si la media aumenta y la concentración también lo hace y de negativa si la media disminuye y la población queda más dispersa aún cuando se debería matizar con relación a la evolución del tamaño de cada categoría. Ninguno de estos casos aparecen en nuestro análisis, la variación entre periodos es del primer tipo salvo en $t_6 - t_7$ que es del segundo tipo. De acuerdo con estos criterios la introducción de las medidas de t_5 , t_6 y t_8 se podría calificar de moderadamente positiva y la introducción de la medida de t_7 , de moderadamente negativa. Por otra parte, en el caso de los presentados, la situación cambia ligeramente. Aparecen un nuevo tipo de comportamiento:

3. μ y *sat* aumentan mientras que σ y *gini* disminuyen.

Este tipo de comportamiento se puede observar tanto en $t_5 - t_6$ como en $t_7 - t_8$. En el primer caso el decremento de σ es casi inapreciable mientras que en el segundo caso es algo mayor. De acuerdo con los criterios propuestos, sobre el conjunto de notas de presentados, la introducción de la medida t_8 se puede calificar de positiva mientras que las medidas t_5 y t_8 se podría calificar de nuevo como moderadamente positivas y la medida t_7 continuaría como moderadamente negativa.

Observemos que el análisis variacional llevado a cabo en el apartado 3 proporciona una clasificación diferencial de los periodos que siguen el primer tipo de comportamiento.

Finalizamos este estudio analizando la evolución de los indicadores μ y *sat*. Para ello hemos calculado los valores de estos dos estadísticos para la población de estudiantes y la de estudiantes presentados desagregados tanto por rango de nota como por segmento (10 %, 40 %, 50 %). En la figura 4 se encuentran los resultados para μ y en la figura 5 los correspondientes a *sat*. Con relación a la desagregación por notas no representamos la categoría *NP* ya que en este caso la nota es cero y ello nos proporciona valor 0 a todos los indicadores. En el caso de la media, observamos que la tendencia de la media en *Est* se preserva tanto en *Apr* como en *NR*, suavizándose en el tramo creciente y con una mayor repercusión negativa en *NR* de la medida introducida en t_7 . El resto de categorías de notas no presentan variaciones significativas de la media y las curvas son casi planas. Desde el punto de vista de segmentos la tendencia en *Est* se presenta en todos ellos algo más suavizada en *Eb*. En el caso de *EstPr* las tendencias no varían significativamente con relación a *Est* observándose un incremento en los valores. Como puede apreciarse en la figura 5, las tendencias de la desagregación por segmentos son más similares en *EstPr* que en *Est*.

Un comportamiento similar aparece con relación a *sat*. En todas las categorías (menos *NR* y *Apr*) las curvas son casi planas. El factor de corrección mantiene la tendencia de *Est* en *Apr* y se modifica significativamente los valores de *sat* en *NR*. En la desagregación por segmentos la tendencia de *Est* y de *EstPr* se mantiene en todos ellos, suavizándose en *Eb* y con pendientes algo más pronunciadas en *Em*.

Cabe observar también las diferencias en el posicionamiento de las curvas desagregadas dependiendo de la medida. La curva de *Rev* está claramente por encima de la de *Est* en la desagregación de *sat* y no tanto en la desagregación de μ . Esta inversión de posicionamiento es aún más clara cuando se compara *Rev* con *EstPr*. En el caso de *NR* y *Sus* la curvas de μ y *sat* invierten su posición relativa. Este último tipo de inversión también se da entre *Est* y el segmento *Em* mientras que no se da cuando se desagrega *EstPr*.

El análisis de los resultados de este estudio parece indicar que *sat* proporciona una mejor predicción de la tendencia de todos los conjuntos desagregados mientras que las tendencias de las curvas correspondientes a μ presentan una mayor variabilidad.

	<i>Est</i>	μ	<i>md</i>	<i>sat</i>	σ	<i>gini</i>
<i>t</i> ₄	514	2,83	2,01	1,29	2,77	0,54
<i>t</i> ₅	492	3,73	3,40	1,97	3,08	0,47
<i>t</i> ₆	466	4,39	4,29	2,49	3,30	0,43
<i>t</i> ₇	436	3,49	3,10	1,78	3,02	0,49
<i>t</i> ₈	448	3,80	3,80	2,04	3,08	0,46

(a) Indicadores sobre matriculados

	<i>EstPr</i>	μ	<i>md</i>	<i>sat</i>	σ	<i>gini</i>
<i>t</i> ₄	440	3,30	2,80	1,76	2,72	0,47
<i>t</i> ₅	445	4,15	4,10	2,44	2,97	0,41
<i>t</i> ₆	395	5,15	5,44	3,44	2,96	0,33
<i>t</i> ₇	343	4,39	4,50	2,81	2,75	0,36
<i>t</i> ₈	342	4,98	5,15	3,51	2,56	0,29

(b) Indicadores sobre presentados

Cuadro 7: Indicadores sobre matriculados y presentados.

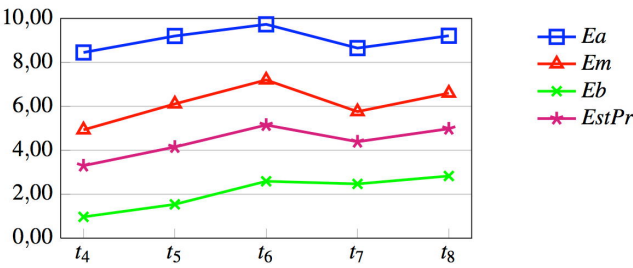
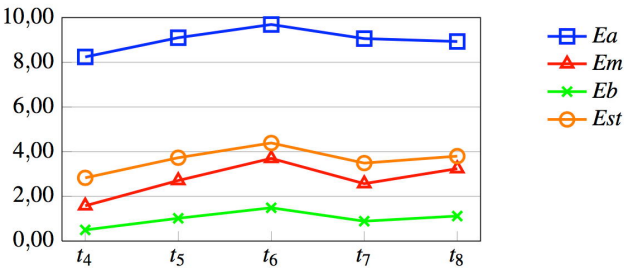
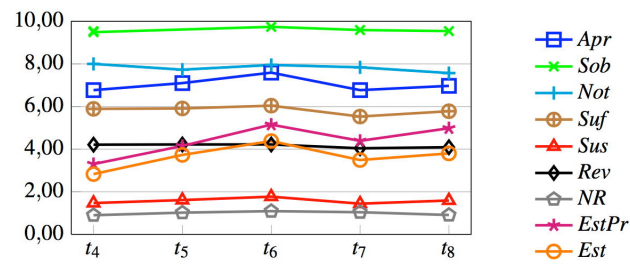
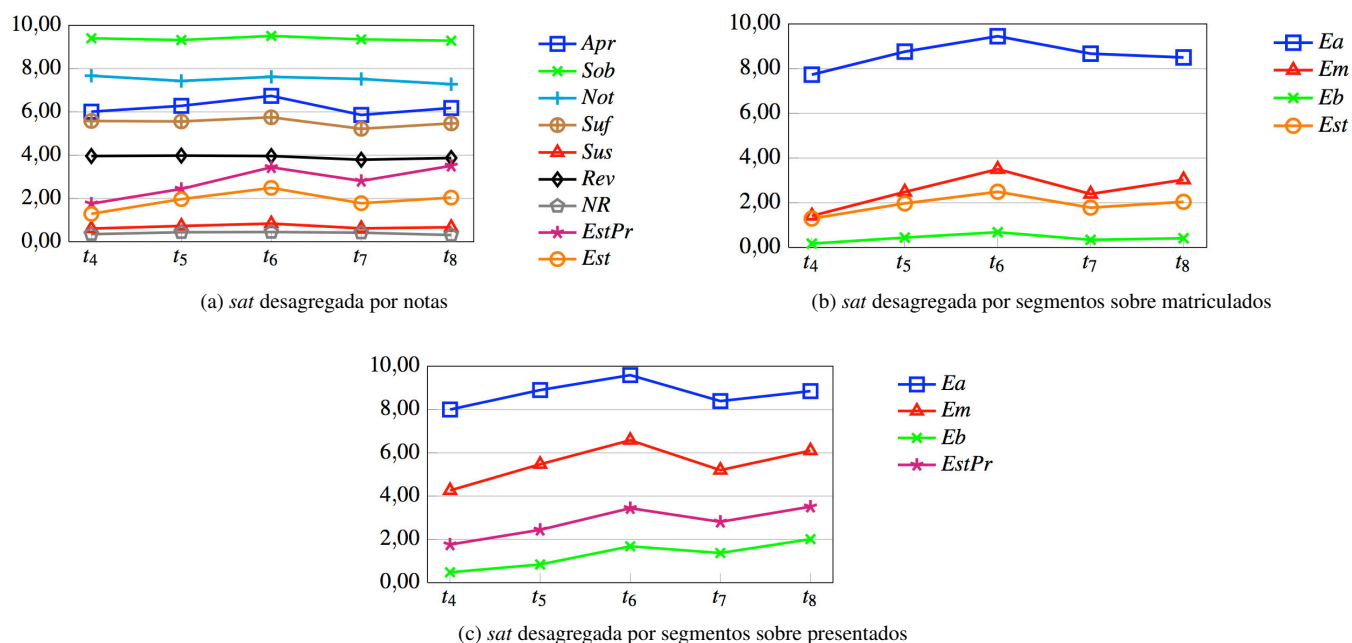


Figura 4: Análisis de la media (μ).

Figura 5: Análisis de la satisfacción (*sat*).

6. Conclusiones y líneas complementarias

Este artículo presenta un análisis de la evolución temporal del curso Programación-1 (PRO1) a lo largo de diferentes semestres utilizando herramientas del área de la economía y de las ciencias sociales. En particular se hace un análisis coste-beneficio de la productividad del curso que se mide en función del trabajo de los docentes y de los resultados de los estudiantes, desagregándolos según diferentes categorías de notas. La distribución de los resultados se interpreta también desagregando los resultados de los estudiantes en el 10 % alto, el 40 % medio y el 50 % bajo. Los resultados se han analizado para los estudiantes matriculados, presentados y aprobados. Las variaciones de estas distribuciones se caracterizan usando el coeficiente de Gini. Finalmente, se analiza la evolución de los indicadores μ (media) y *sat* (satisfacción), desagregando tanto por rango de nota como por segmento (10 %, 40 %, 50 %).

El análisis del coste-beneficio desagregado por categorías de notas expande el realizado anteriormente [4] mostrando que el efecto de la productividad se extiende a todas las categorías y no se reparte uniformemente. Además, la ganancia marginal desagregada permite medir con más detalle la bondad de las diferentes medidas introducidas en el curso sobre de las diferentes poblaciones. Todo ello refuerza la tesis de que el estudio coste-beneficio de la productividad aumenta la comprensión de la evolución de los cursos y, por lo tanto, proporciona elementos para diseñar futuras estrategias. El análisis de la desigualdad en los resultados muestra que a lo largo de los años el esfuerzo destinado a la mitad de nota más baja ha sido

infructuoso. Por último, la evolución de μ y *sat* es muy similar tanto por notas como por segmentos, lo que indica que el factor de corrección aplicado por *sat* no cambia la tendencia de μ .

Uno de los aspectos a tener en cuenta en un futuro estudio es la incorporación de un tratamiento diferenciado de los estudiantes con *NP*. Este tratamiento diferenciado requerirá de un análisis detallado de los plazos y períodos en los que se produce el abandono. Esto permitiría una estimación más ajustada de la carga de trabajo por estudiante y de la carga total de los docentes. La tarea no es fácil, pero una aproximación sería factible teniendo en cuenta, además de las notas finales, las notas de todos los actos evaluatorios. Podría ser interesante estudiar las discrepancias y coincidencias entre σ y *gini* en nuestro caso, para ello se podrían utilizar las propuestas de Atkinson [2].

Otros aspecto de la implantación de cada medida que no han sido considerados en este estudio debido a la ausencia de datos fiables son, por una parte, la cantidad de trabajo dedicado por los estudiantes a la asignatura y, por otra, la consideración de cómo cada una de las medidas introducidas es percibida por los profesores y los estudiantes.

Una aproximación de la cantidad de trabajo realizado por los estudiantes se podría extraer del registro de envíos al Jutge, sin embargo en este caso se requieren estudios adicionales para poder determinar cuáles son los parámetros que reflejan mejor esta dedicación. Por poner un ejemplo, el número de problemas resueltos satisfactoriamente parece un posible indicador, pero quizá el número de problemas que se han intentado resolver o cuánto esfuerzo (en tiempo, en intentos, etc.) se ha puesto en cada uno de ellos es un parámetro mejor. Desarro-

llar herramientas para extraer información adicional del Jutge que permitan analizar el volumen de trabajo dedicado en un período de tiempo es una dirección muy interesante para futuros estudios. La segunda dirección permitiría tomar en consideración aspectos totalmente subjetivos de la implantación de cada medida. Puede suceder que medidas para las que “objetivamente” alguno de los indicadores no tiene una evolución positiva se perciban como positivas por alguno de los colectivos o viceversa. Por tanto, creemos que puede resultar interesante complementar el estudio propuesto en este trabajo con un análisis de la percepción de los distintos colectivos implicados. Desafortunadamente, no disponemos de los datos que permitirían realizar dicho estudio. Para ello, tendríamos que disponer de encuestas diseñadas de forma específica y ad-hoc para cada uno de los colectivos. La UPC realiza regularmente encuestas¹ a sus estudiantes con el fin de tener indicadores del funcionamiento de las asignaturas así como del grado de satisfacción de los propios estudiantes. Sin embargo, la encuesta es uniforme para todas las asignaturas pidiendo opiniones sobre los aspectos genéricos del desarrollo de una asignatura: interés del contenido, adecuación de la evaluación, coordinación de las distintas actividades y nivel de satisfacción. Por otra parte el nivel de participación (por debajo del 30% en las encuestas del período comprendido en este estudio) es demasiado bajo como para permitir un análisis fiable. Pensamos que se debería conseguir realizar encuestas que proporcionen información sobre el desarrollo de la asignatura (aunque quizá de modo indirecto) con alta participación. Es importante tener indicaciones sobre las expectativas de los diferentes colectivos implicados [9, 22], lo que les motiva y satisface [15], y cuantificar su grado de satisfacción. Estudios en este sentido del desarrollo de la actividad docente serían deseables y permitirían complementar el análisis basado en datos objetivos.

Agradecimientos

Joaquín Gabarró agradece a Jerto Sempere, Albert García y a Enric Bech las largas discusiones sobre renta y desigualdad. Los autores agradecemos los útiles comentarios del revisor, que sin duda ayudaron a mejorar el contenido y la presentación de este manuscrito. Trabajo parcialmente financiado por la mención de la Generalitat de Catalunya a ALB-COM como grupo de investigación consolidado (ref. SGR 2014:1034) y por los proyectos TIN2013-46181-C2-1-R y TIN2012-37930-C02-02 del Ministerio de Economía y Competitividad y fondos FEDER de la Unión Europea.

Referencias

- [1] Kenneth J. Arrow. The economic implications of learning by doing. *The Review of Economic Studies*, vol. 29, núm. 3, pp. 155–173, 1962.

¹<https://e-enquestes.upc.edu>

- [2] Anthony B. Atkinson. *On the measurement of inequality*. *Journal of Economic Theory*, vol. 2, núm. 3, pp. 244–263. 1970.
- [3] Anthony B. Atkinson. *Desigualdad. ¿Qué podemos hacer?* Fondo de cultura económica. México. 2016.
- [4] Maria J. Blesa, Amalia Duch, Joaquim Gabarró, Jordi Petit y Maria Serna. *Continuous assessment in the evolution of a CS1 course: the pass rate/workload ratio*. En Susan Zvacek, Maria Teresa restivo, James Uhomobhi y Markus Helfert (Editores): *Computer Supported Education*. Serie Communications in Computer and Information Science, vol. 583, pp. 313–332. Springer. 2016.
- [5] Ludovic Bot, Pol-Bernard Gossiaux, Carl-Philippe Rauch y Safouana Tabiou. *Learning by doing: a teaching method for active learning in scientific graduate education*. *European Journal of Engineering Education*, vol. 30, núm. 1, pp. 105–119. Marzo de 2005.
- [6] Samuel Bowles. *Towards an educational production function*. En W. Lee Hansen (editor): *Education, Income, and Human Capital*, pp. 11–70. National Bureau of Economic Research. 1970.
- [7] Lidia Ceriani y Paolo Verme. *The origins of the Gini index: extracts from Variabilità e Mutabilità (1912) by Corrado Gini*. *The Journal of Economic Inequality*, vol. 10, núm. 3, pp. 421–443, 2012.
- [8] Denis Diderot. *Arithmétique politique*. Artículo de L’Encyclopédie. 1751.
- [9] Albert García Pujol. *Integrar nous conceptes al PIB augmenta la seva importància*. *Informatiu de l’economista*, núm. 155, pp. 44–46. Marzo de 2016.
- [10] O. Giménez, J. Petit y S. Roura. *Programació 1: A pure problem-oriented approach for a CS1 course*. En actas del Informatics Education Europe IV (IEE), pp. 185–192. 2009.
- [11] Omer Giménez, Jordi Petit y Salvador Roura. *Jutge.org: an educational programming judge*. En actas del 43rd ACM Technical Symposium on Computer Science Education (SIGCSE’12), pp. 445–450. Raleigh, Carolina de Norte, USA. 2012.
- [12] Corrado Gini. *Variabilità e Mutabilità. Contributo allo Studio delle Distribuzioni e delle Relazioni Statistiche*. Paolo Cuppini. Bolonia. 1912.
- [13] Eric A. Hanushek. *Education production functions*. En *The New Palgrave Dictionary of Economics*. Palgrave Macmillan. 2008.

- [14] Petri Ihantola, Tuukka Ahoniemi, Ville Karavirtay y Otto Seppälä. *Review of recent systems for automatic assessment of programming assignments*. En actas del 10th Koli Calling International Conference on Computing Education Research, pp. 86–93. Koli, Finlandia. 2010.
- [15] Richard Layard. *Happiness: Lessons from a New Science*. Penguin. 2011.
- [16] F.J. Martín-Carrasco, A. Granados, D. Santillán y L. Mediero. *Continuous assessment in civil engineering education. Yes, but with some conditions*. En actas del 6th International Conference on Computer Supported Education (CSEDU), vol. 2, pp. 103–109. Barcelona. Abril de 2014.
- [17] Thomas Piketty. *El capital en el siglo XXI*. RBA. 2015.
- [18] Miguel A. Revilla, Shahriar Manzoor y Rujia Liu. *Competitive learning in informatics: The UVa online judge experience*. Olympiads in Informatics, vol. 2, pp. 131–148. 2008.
- [19] James Robergé, Matthew Bauer y George K. Smith. *Engaged learning for programming in C++: A laboratory course. 2nd edition* Jones and Bartlett. Sudbury, Massachusetts, USA. 2001.
- [20] Amartya Sen. *Real national income*. The Review of Economic Studies, vol. 43, núm. 1, pp. 19–39. 1976.
- [21] Robert M. Solow. *Learning from ‘Learning by Doing’ Lessons for Economic Growth*. Stanford University Press. 1997.
- [22] Joseph E. Stiglitz, Amartya Sen y Jean-Paul Fitoussi. *Medir nuestras vidas: Las limitaciones del PIB como indicador de progreso*. RBA. 2013.
- [23] Joseph E. Stiglitz y Bruce C. Greenwald. *Creating a learning society: A new approach to growth, development, and social progress*. Columbia University Press. 2014.
- [24] Vinod Thomas, Yan Wang y Xibo Fan. *Measuring education inequality: Gini coefficients of education*. World Bank research working paper; núm. WPS 2525. Washington, DC. 2001. Disponible en <http://documents.worldbank.org/curated/en/361761468761690314/Measuring-education-inequality-Gini-coefficients-of-education>. 2001.
- [25] Vinod Thomas, Yan Wang y Xibo Fan. *Measuring education inequality: Gini coefficients of education for 140 countries (1960–2000)*. Journal of Education Planning and Administration, vol. 17, núm. 1, pp. 5–33. 2003.
- [26] Neilor A. Tonin, Fabio A. Zanin y Jean Luca Bez. *Enhancing traditional algorithms classes using URI online judge*. En actas del 2012 International Conference on e-Learning and e-Technologies in Education (ICEEE), pp. 110–113. Septiembre de 2012.
- [27] Hal R. Varian. *Intermediate Microeconomics: A Modern Approach*. W. W. Norton and Company, 7ª edición. 2005.
- [28] Elena Verdú, Luisa M. Regueras, María J. Verdú, José P. Leal, Juan P. de Castro y Ricardo Queirós. *A distributed system for learning programming on-line*. Computers & Education, vol. 58, núm. 1, pp. 1–10. 2012.



Maria J. Blesa obtuvo en 2006 el título de Doctora Ingeniera en Informática por la Universitat Politècnica de Catalunya, donde es Profesora Agregada del Departamento de Ciències de la Computació desde el año 2007. Imparte clases en la Facultat de Informàtica de Barcelona y en la Facultat de Matemàtiques i Estadística. Sus áreas de investigación principales incluyen la algoritmia, la teoría algorítmica de juegos y la optimización combinatoria. Su producción científica supera las 50 contribuciones en revistas internacionales, congresos y capítulos de libro.



Amalia Duch realizó sus estudios de doctorado en el departamento de Ciencias de la Computación de la Universitat Politècnica de Catalunya donde es Profesora Agregada desde el año 2005. Imparte cursos de Programación y Estructuras de Datos y Algoritmos en la Facultat de Informàtica de Barcelona y en su trabajo de investigación se interesa principalmente por temas de estructuras de datos multidimensionales y análisis de algoritmos.



Joaquim Gabarró, nació en Barcelona en 1953 y estudió Física e Informática. Obtuvo el Doctorado en la Universitat Politècnica de Catalunya (UPC) en 1983. Desde 1985 es Profesor Titular de Universidad en el Departamento de Ciencias de la Computación de la UPC. Es coautor en más de 70 contribuciones científicas en revistas internacionales, actas de congresos y capítulos de libros. Ha sido coautor de dos

libros. Sus principales líneas de investigación son concurrencia, algoritmos y complejidad.



Jordi Petit es el desarrollador principal de Jutge.org.

Jordi Petit obtuvo el título de Doctor Ingeniero en Informática por la Universitat Politècnica de Catalunya en 2001. Desde 2002 es Profesor Titular de la Universitat Politècnica de Catalunya, trabajando en el departamento de Ciències de la Computació. Sus intereses actuales de investigación son la ingeniería de algoritmos y los juegos educativos de programación.



Maria Serna es Licenciada en Matemáticas por la UB (1981), Licenciada en Informática (1985) y Doctora en Informática (1990) por la UPC. En 1982 se incorporó como docente a la UPC como miembro del Departamento de Matemática Aplicada I hasta el año 1992 en el que se incorporó al actual departameto de Ciencias de la Computación. Desde 2006 es Catedrática de

Universidad y coordinadora del Programa de Doctorado en Computación. En los últimos años su trabajo docente se ha centrado en diversas asignaturas de algorítmia y programación. Ha dirigido/participado en más 25 proyectos de investigación, 11 de ellos europeos, y participado como miembro o Presidente de Comité de Programa de unos 50 congresos. Ha sido adjunta en el equipo de coordinación del panel Ciencias de la Computación y Tecnologías Informáticas (TIN) de la Agencia Nacional de Evaluación y Prospectiva y participado como evaluador externo en diversas convocatorias de la CE. Es coautora de más de 120 contribuciones científicas en revistas internacionales, actas de congresos y capítulos de libros. Sus líneas de investigación principales son algorítmia y complejidad, computación distribuida y paralela, problemas computacionales en grafos y redes junto con la teoría algorítmica de juegos.



2017 M.J. Blesa, A. Duch, J. Gabarró, J. Petit y M. Serna. Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional que permite copiar, distribuir y comunicar públicamente la obra en cualquier medio, sólido o electrónico, siempre que se acrediten a los autores y fuentes originales y no se haga un uso comercial.